The Globus File Yanker:
A Remote Web Service for Third Party Grid File Transfers

Shreyas Cholia (NERSC)
Adam Megacz (XWT Foundation)
Stephen Chan (NERSC)

The Globus File Yanker offers a convenient, lightweight Web interface to filesystems served by GridFTP and Grid-enabled HPSS FTP servers, and provides the ability to directly perform a third party copy between these servers. The client uses XWT (XML Windowing Toolkit) to render the user interface. This means that the user can log into the system from a standard Web browser as long as it supports Java or ActiveX.

Components:

1. **XWT CLIENT** This is a Java- or ActiveX-based Web client that is loaded using your Web browser. It consists of a client GUI engine that renders an XML specified UI. It is fairly lightweight, and is only responsible for submitting XML-RPC-based requests to a Tomcat Server and rendering the responses from those requests. Refer to http://www.xwt.org for details.

2. **TOMCAT GRIDFTP PORTAL:** The Apache Tomcat Server communicates with the XWT client using XML-RPC. The client sends the Portal a request, along with authentication information. Each request is serviced by a Java Servlet within Tomcat. The Portal is then responsible for acquiring credentials from the MyProxy Server and talking to the appropriate GridFTP Servers. The Servlets use the Java COG kit to handle the various Globus calls. The Tomcat Portal is also responsible for keeping track of file transfer jobs submitted and the status of those jobs.

3. **MYPROXY SERVER:** The MyProxy Server holds the users' credentials, and delegates a proxy certificate to the Tomcat Portal when presented with a valid username–password combo.

4. **HPSS GRIDFTP SERVER:** This is an HPSS FTP Server with special modifications to support GSI authentication and some of the Grid extensions to the FTP protocol. This server can access files stored in an HPSS storage system.

5. **GRIDFTP SERVER:** This is the standard GridFTP server that is distributed with Globus 2.0. This server supports the Grid extensions to the FTP protocol and serves a standard UNIX file system.

6. **HPSS:** The High Performance Storage System is the mass storage archive.

7. **THIRD PARTY COPY:** This represents a direct file transfer between two Grid-enabled FTP servers. The data flows directly from server to server without going through the client.

How it Works:

The user brings up a Web browser and downloads the XWT client engine, which is run as a Java Applet. He then downloads the XML client interface, which is rendered by the engine.

The client uses XML-RPC to talk to the Tomcat GridFTP Portal Server. A client request is accompanied by a username and password combination. This combination is used by the Portal to acquire a credential for that user from the MyProxy Server.
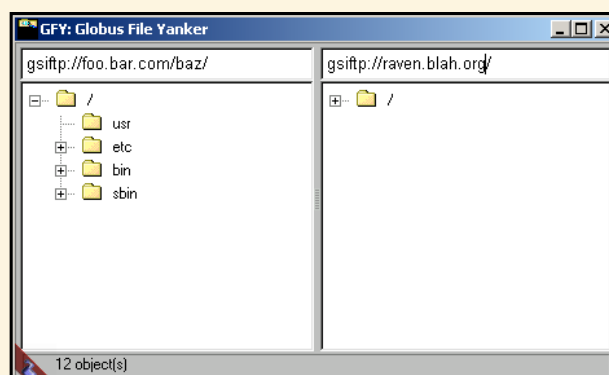
This credential is then used by the Portal to authenticate to the GridFTP Servers. Based on the client request, the Portal will then submit a file transfer job. A file transfer thread will service this request and issue a third party copy between the two servers. Once the job has been submitted the client can query the portal for the status of a job.

Note: Special hooks have to be put into the Portal Servlet code/Java COG kit and in the HPSS FTP Server code, in order to support this service. Currently, this service does not support striped transfers.

Links:
XWT: http://www.xwt.org
Globus: http://www.globus.org
HPSS: http://www.clearlake.ibm.com/hpss
Apache Tomcat: http://jakarta.apache.org/tomcat/

Remote Web Service for third party Grid file transfers